**Marine Genomics**
September 2011, Volume 4, Issue 3, Pages 181-188
http://dx.doi.org/10.1016/j.margen.2011.04.003




# Expressed sequence tags from heat-shocked seagrass *Zostera noltii* (Hornemann) from its southern distribution range

Sónia I. Massa[a], Gareth A. Pearson[a], Tânia Aires[a], Michael Kube[b], Jeanine L. Olsen[c], Richard Reinhardt[b], Ester A. Serrão[a], Sophie Arnaud-Haond[d, *]

[a] CCMAR-CIMAR, Universidade do Algarve, Gambelas; 8005–139 Faro; Portugal
[b] Max-Planck Institute for Molecular Genetics, Ihnestraße 63/73; 14195 Berlin; Germany
[c] Department of Marine Benthic Ecology and Evolution, Centre for Ecological and Evolutionary Studies, Biological Centre, University of Groningen, 9750 AA Haren; The Netherlands
[d] IFREMER, Centre de Brest, BP 70, 29280 Plouzané; France

*: Corresponding author : Sophie Arnaud-Haond : Tel.: + 33 2 98 22 43 05; fax: + 33 2 98 22 47 57 ; email address : Sophie.Arnaud@ifremer.fr

**Abstract :**

Predicted global climate change threatens the distributional ranges of species worldwide. We identified genes expressed in the intertidal seagrass *Zostera noltii* during recovery from a simulated low tide heat-shock exposure. Five Expressed Sequence Tag (EST) libraries were compared, corresponding to four recovery times following sub-lethal temperature stress, and a non-stressed control. We sequenced and analyzed 7009 sequence reads from 30 min, 2 h, 4 h and 24 h after the beginning of the heat-shock (AHS), and 1585 from the control library, for a total of 8594 sequence reads. Among 51 Tentative UniGenes (TUGs) exhibiting significantly different expression between libraries, 19 (37.3%) were identified as 'molecular chaperones' and were over-expressed following heat-shock, while 12 (23.5%) were 'photosynthesis TUGs' generally under-expressed in heat-shocked plants. A time course analysis of expression showed a rapid increase in expression of the molecular chaperone class, most of which were heat-shock proteins; which increased from 2 sequence reads in the control library to almost 230 in the 30 min AHS library, followed by a slow decrease during further recovery. In contrast, 'photosynthesis TUGs' were under-expressed 30 min AHS compared with the control library, and declined progressively with recovery time in the stress libraries, with a total of 29 sequence reads 24 h AHS, compared with 125 in the control. A total of 4734 TUGs were screened for EST-Single Sequence Repeats (EST-SSRs) and 86 microsatellites were identified.

**Highlights**

► Response to heat stress is very fast but gene expression returns to normal after 24 h. ► Photosynthesis-related genes were under-expressed after heat-shock. ► Heat-shock caused a quick rise in heat shock proteins and molecular chaperone expression.

**Keywords :** EST library ; High temperature stress ; Abiotic stress response ; Climate change

## 1. Introduction

Seagrasses are key-species in coastal ecosystems, as important primary producers providing food, nursery and shelter for many marine species, but are declining worldwide [1, 2]; changes attributed to anthropogenic activities and climate change [1, 3]. The dwarf seagrass *Zostera noltii* dominates the intertidal habitats of the coastal lagoon system Ria Formosa, in southern Portugal, where at low tide this species occurs either in shallow intertidal pools or completely emersed. The predicted increase in sea surface temperatures (SST) of 0.2ºC/decade [4] is raising concerns as to the ability of this species to survive the expected global warming, particularly since intertidal habitats are more affected by air temperature than SST. Previous work suggested that the current temperature in shallow intertidal pools is already very close to the temperature tolerance of approximately 38ºC determined for *Z. noltii* in the Ria Formosa [5], with records showing temperatures as high as 36ºC in the summer of 2007. Plants that occur in the upper intertidal can be exposed for as long as 6 hours during spring tides, and temperature in these small intertidal pools have reached 38ºC in low wind/high sun and air temperature situations (S. Massa, personal observation).

When higher plants are exposed to temperatures higher than required for optimal growth, a cellular adaptive response is activated to maintain cellular homeostasis under stress, resulting in increased synthesis of heat-shock proteins (HSP) and reduced production of other metabolic proteins [6]. Heat-shock can affect macromolecular synthesis, levels of cations, states of protein phosphorylation, metabolic pathways and cytoskeleton networks [7]. Heat stress has been shown to reduce photosynthetic yield in tropical seagrasses and increase photoinhibition [8]. It also reduces chlorophyll levels, reducing light (energy) absorption in chloroplasts for photosynthesis, and affects the levels of protective enzymes against oxidative stress, such as ascorbate peroxidase

(APX) and superoxide dismutase (SOD) [9, 10]. All of this suggests that the response to heat stress is a complex process involving a large number of genes.

As genomic tools become more accessible for a wide range of non-model organisms, they can be applied by evolutionary ecologists to address questions concerning adaptation to novel or stressful environments, in order to clarify whether changes in the phenotype of independent populations confronted with similar environmental challenges are adaptive or a result of random genetic drift [11-13]. Differences in individual responses to adverse conditions affect fitness through selection, and may ultimately influence the survival of populations and potential shifts in species distribution ranges [14].

As of February 5$^{th}$, 2011 there were 13,434 marine seagrass EST sequences reads in NCBI, from only 2 species, of which 10,345 are from the closely related species *Zostera marina* and 3,089 from *Posidonia oceanica*, and none for *Z. noltii.* This study aims to identify genetic markers and genes expressed in the intertidal seagrass *Z. noltii* during recovery from heat-shock exposure, simulating a low-tide event. A description of transcriptional changes in *Z. noltii* in response to heat-shock is needed to study the impacts of temperature stress and to explore the adaptation potential of this species in the face of climate change. (e.g. molecular markers including SSR and SNPs to identify loci potentially under selection, and gene sequence information to develop qPCR assays for functional studies of target genes). Moreover, an increased coverage of the transcriptome of seagrasses in general will allow the identification of crucial clusters of genes systematically involved in response to warming, providing a basis for comparative studies of gene evolution in seagrasses.

## 2. Material and methods

### *2.1. Culture conditions and stress treatment*

*Z. noltii* plants were collected in the spring of 2007 at 4 sites in the Ria Formosa (Ramalhete, Praia de Faro, Olhão and Portimão) in several distant cores with natural sediment, and placed in an outdoor tank (1.5 x 1.5 m, maximum seawater volume 225 L) with ambient light conditions, continuous seawater flow from the Ria Formosa and simulated tides for acclimation during approx. 4 weeks. In high tide situation, water level in the tank was up to its maximum capacity, and every day at 10 a.m. the water level was slowly decreased until roughly two hours later only a 2 cm layer of water (approximately) remained above the sediment, mimicking a low tide situation. At 4 p.m., ambient seawater was gradually added to increase the water level back to high tide situation. This process was repeated again starting at 10 p.m. as tides in the Ria Formosa are semidiurnal. Previous work had determined the sub-lethal temperature for *Z. noltii* shoots in Ria Formosa to be slightly above 37ºC [5], and so a single heat-shock of 37.5±0.5 ºC was then applied for four hours, the approximate average duration of low tide exposure in this part of *Z.noltii* distributional range, during a simulated low tide situation as previously described, between 10 am and 2 pm. After the heat-shock, ambient seawater was gradually added to the tank to lower the temperature to its initial value of approx. 22ºC. Sampling of inner leaves of randomly picked shoots (on average, 59 x approx. 20 shoots per time step) from all sites and cores for RNA extraction occurred at 30 minutes, 2, 4 and 24 hours after the beginning of the heat-shock. Each group of approximately 20 shoots was put in a tube and immediately frozen in liquid nitrogen for later processing. A sample of non-stressed plants was also collected to be used as a control (blank).

### *2.2. RNA preservation and storage*

At the end of the first three samplings (plus the control), samples were taken to the lab in liquid nitrogen and immediately freeze-dried for at least 48 hours, and preserved at –80ºC until extraction. The same procedure was performed after the 24 hour sampling.

### *2.3. cDNA library construction*

Samples were transferred into a tube with a tungsten sphere and ground at 30 g for 10 minutes. RNA extraction was performed using Qiagen and GE Healthcare extraction kits (Germany). RNA quality was verified using 3µL of RNA extraction on denaturing agarose gels and quantification was performed with a spectrophotometer at 260 and 280 nm.

RNA extractions were treated with the Macherey-Nagel NucleoSpin RNAII kit for DNAse I digestion. All cDNA libraries were constructed with purified mRNA using Dynabeads® mRNA purification kit (Invitrogen); each sample was fractionated by column chromatography and the highest quality fractions were pooled together and directionally cloned into pDORN 222, using CloneMiner™ cDNA Library Construction Kit (Invitrogen). Electrocompetent cells were transformed by electroporation and then grown in SOC medium. After plasmid extraction, sequence data were obtained using M13-21 primer (5'-TGTAAAACGACGGCCACT-3´), amplified by the following programme: denaturation step (96ºC 1'), followed by 35 cycles of denaturation (96ºC 20''), annealing (55ºC 10'') and elongation (60ºC 4'). The cloned cDNAs were 5′-end-sequenced using Big Dye 3.1 chemistry and ABI 3130XL capillary sequencers.

### *2.4. Sequence processing and EST assembling*

Each library was treated independently to allow comparison of the response to stress over time. An average of 1,719 clones were sequenced for each of the five treatments

(30 min, 2h, 4h, 24h and control) and analysed with the ESTragon program (developed and routinely used in the Max Planck Institute for Molecular Genetics) for removal of low quality [15] and/or contaminated sequence reads from each dataset before alignment, with a cut-off length of 100 bp. All vector-clipped and high quality sequence reads were submitted to dbEST within GenBank [accession numbers HO214335-HO215643 for the control library, HO215644-HO217278 for 30 minutes, HO217279-HO218949 for 2 hours, HO218950-HO220485 for 4 hours and HO220486-HO222020 for 24 hours]. Processed sequence reads were then assembled into contigs to cluster the individual sequence reads to represent unique transcripts using TGICL [16]. Unique sequences, either singletons or contigs, will from now on be referred to as Tentative UniGenes (TUGs). Blastx analysis was performed with Blast2GO® (V 2.4.4) [17, 18] against NCBI non redundant (nr) protein database, using a preliminary E-value of $10^{-3}$ as cutoff, to maximise the number of annotation terms obtained. Blastx results were then mapped and annotated against the Gene Ontology (GO) database, according to three different categories: Biological Process (BP), Molecular Function (MF) and Cellular Component (CC). Annotation parameters were as follows: E-value Hit filter $1.0E^{-6}$, annotation cutoff 55, GO Weight 5. Fisher's Exact Tests were performed to test for significance of differential expression in each heat-shock library *versus* the control library.

### *2.5. Homology searching*

Individual stress libraries were screened by local Blastn against a database consisting of a pooled assembly of all stress treatment libraries (i.e., 30 min, 2 h, 4 h stress, and 24 h recovery) to identify homologues and estimate their expression levels. In this case, when two or more TUGs from individual stress libraries had the same top blast hit

against this database, they were considered the same TUG. In the case of the control library that was not included in the database, we applied a stringent E-value cut-off of < $1.0E^{-100}$ when considering homologous consensus TUG sequences.

TUGs with significantly different expression levels over the time-course of the stress treatment, or between stressed and non-stressed samples, were identified with the IDEG6 (Identification of Differentially Expressed Genes) web tool [19], using the Audic and Claverie method [20] to compare, after normalization, the number of sequence reads/TUG in each library (significant threshold of $\alpha = 0.05$, corrected for multiple testing with [(m+1)/2m]). Only TUGs with a total of at least 10 sequence reads from all libraries were used in the expression analysis, in an attempt to control for random sequencing effects in small TUGs. The analysis resulted in a total of 51 TUGs with significantly different expression in at least one pairwise library comparison. These were then screened using Blastn against Dr. Zompo's EST databases for *Z. marina* and *P. oceanica* to find interspecific matches [21, 22] with an bit score cutoff value of 200.

### *2.6. Microsatellite identification*

All TUGs from all 5 libraries were screened for single sequence repeats (EST-SSRs) or microsatellites using MsatCommander [23], for all motifs with two to six nucleotides length and a minimum of six repeat units.

## 3. Results and Discussion

### *3.1. EST sequencing determination*

A total of 8,594 clones from all libraries were sequenced, resulting in 7,799 high quality sequence reads. After screening for cloning relics and contamination, a total of 113 reads were removed, leaving a total of 7,686 successful reads. Clustering (TGICL)

resulted in 170 contigs for the 30 min library, 174 contigs for the 2h library, 151 contigs for the 4h library, and 145 contigs for both the 24h and the control libraries. Average percentage of sequence reads with a successful blast against NCBI non redundant (nr) protein database was 82%, of successfully mapped sequences 70% and successfully annotated against the GO database was 56%. Detailed information for each library can be found in table 1.

The main protein families found in our data were the *HSP70* protein family, the protein kinase domain family and the chlorophyll a-b binding protein family (Table 2). The reduction in the number of sequence reads of the *HSP70* protein family between the stress and the control libraries was very clear, while the number of chlorophyll a-b binding proteins was roughly the same among libraries.

### *3.2. Identification of differential expression*

Out of the 51 TUGs with significantly different expression between libraries and a total of at least 10 sequence reads, 26 (51%) had a matching hit with *Z. marina*, 16 (31%) of which also had a match with *P. oceanica* (Annex I). Most of the highest-scoring matches were photosynthesis-related TUGs but also included 2 heat-shock cognate protein 80 TUGs, a stress-induced phosphoprotein, a glyceraldehyde-3-phosphate dehydrogenase, a flavoprotein, a 70 kDa peptidyl-prolyl isomerase, a glutamine synthetase and one TUG with no significant Blastx hit, suggesting that these TUGs are highly conserved across the seagrass species, probably due to their fundamental role in cellular pathways. The remaining 13 TUGs (25%) had no meaningful matches. These 51 TUGs were divided into 3 groups according to their function: ‘molecular chaperones’, ‘photosynthesis TUGs’ and ‘other TUGs of interest’, which will be discussed further below.

220

221 ***3.3. Molecular chaperones***

222 Among 943 TUGs from the control library, only two heat-shock protein TUGs were
223 found, and 61 heat-shock protein TUGs were found in the combined stress libraries,
224 from a total of 4,666 TUGs. When the number of sequence reads was taken into
225 consideration, IDEG found 9 of these heat-shock protein TUGs to have a significantly
226 different expression between libraries. 'Molecular chaperones' include not only these 9
227 heat-shock protein TUGs but also 10 other TUGs, like heat-shock cognate proteins and
228 DnaJ homologues, and were over-expressed in the stress libraries. While the control
229 library showed only 2 sequence reads, these increased rapidly at the beginning of the
230 stress, showing a total of 229 sequence reads at 30 minutes, and slowly decreased
231 thereafter to 213 at 2 hours, 126 at 4 hours and 66 at 24 hours (Fig. 1). The response
232 was very fast, with significant difference being recorded as soon as 30 minutes after the
233 beginning of the heat-shock.

234 A total of eight sequences (5 contigs and 3 singletons) encoding small heat
235 shock proteins (*sHSPs*) were identified in our dataset based on pfam database searches
236 of predicted ORFs. ORF prediction was performed using the OrfPredictor tool [24]
237 available at http://proteomics.ysu.edu/tools/OrfPredictor.html. The resulting ORFs were
238 used to identify Pfam domains on the UFO webserver (http://ufo.gobics.de/submission,
239 [25]). All 8 sequences had start sites and encoded a predicted full-length *sHSP* with an
240 *HSP20*/alpha crystalline domain (PF00011). In two separate cases a singleton and
241 contig were found to have identical protein sequences, but variable 3' UTRs. The other
242 3 contigs and one singleton encoded unique proteins. Searches for signal peptides
243 targeting the proteins to the chloroplast or mitochondrion were negative, and no
244 transmembrane domains were identified, suggesting that these proteins are putatively

245 cytoplastic. Heat-shock proteins are highly conserved among species; their main role is
246 to protect and repair protein structures, preventing the denaturation process or
247 promoting the proper refolding of denatured proteins [7]. Since protein conformation is
248 also important in normal conditions, most HSPs are present in all cells and tissues even
249 in the absence of stress factors such as elevated temperatures [26]. Heat-shock proteins
250 are also involved in the apoptotic pathways, as irreparably damaged cells will
251 eventually be eliminated when stress conditions are too severe [27]. Previous work has
252 shown that mortality will eventually occur (Massa et al, unpubl.). After 24 hours the
253 response appeared to be complete as Fisher's Exact Test showed no further differences
254 between 24h expression and the control library (Annex 2), suggesting a return to normal
255 gene expression 24 hours after the heat-shock.

256 The other 10 TUGs categorized as 'molecular chaperones' include DnaJ-like
257 proteins, which stimulate the ATPase activity of *HSP70* protein [28], that catalyse the
258 folding of proteins and the assembly of protein complexes [29, 30] and heat-shock
259 cognate (Hsc) 70 and 80 (Fig 1). Although HS cognates are constitutively expressed, it
260 has been shown that *HSC70* is required for Heat-Shock Factor 1 to become activated
261 and target expression of appropriate genes during heat stress [31]. Molecular
262 chaperones are important during both unstressed and stressful conditions, as they not
263 only assist in the *de-novo* folding of denatured proteins but also in conformational
264 changes that affect function, while they also transport unfolded proteins across
265 membranes and plasmodesmata [32, 33]. The stress-induced phosphoprotein 1, that we
266 also included in this group of 'molecular chaperones', was first described in
267 *Saccharomyces cerevisiae* where it was implicated in mediating the heat-shock response
268 of some *HSP70* genes [34].

269 Our results are in contrast with findings for the closely related species *Z. marina*

[21] where no over-expression of heat-shock proteins was observed in the EST library constructed for elevated growth temperature. However, this may reflect the differences in the organismal biology and ecology of this species in comparison with intertidal *Z. noltii* (the *Z. marina* populations used were subtidal and therefore exist in an environment with considerably lower amplitudes and extreme values of temperature fluctuation). The major aim of the *Z. marina* study was to investigate sub-lethal growth temperature (25ºC; 2ºC higher than an expected summer maximum) rather than heat shock *per se*. If a minor heat-shock response was induced, it may have subsided by the time of sampling, as our results show that the response may be over between 4 and 24 h after the end of the shock.

Reusch and colleagues suggested that the lack of heat-shock protein expression in *Z. marina* might be explained by the fact that permanently submerged seagrasses, as was the case with *Z. marina* in their experiments, are seldom subject to rapid temperature fluctuations because of the 'buffered' aquatic environment, lending contrast to terrestrial environments where the presence of heat-shock proteins genes is a common response to temperature stress [21]. Our results demonstrate that for intertidal marine angiosperms like *Z. noltii*, the geographic distribution of which extends to warmer latitudes where *Z. marina* is strictly subtidal, temperature in small intertidal pools is indeed more influenced by air than by sea temperatures; the former reach levels very close to the sub-lethal temperature of *Z. noltii* shoots during summer in the Ria Formosa [5]. The regular and rapid increase of temperature in small intertidal pools, together with potential desiccation, may therefore require the maintenance of a heat stress response in dwarf eelgrass more similar to terrestrial angiosperms than subtidal ones.

### *3.4. Photosynthesis TUGs*

We found 12 TUGs with photosynthetic function and with significantly different expression between libraries, which were overall under-expressed in the stress libraries. The control library showed a total of 125 sequence reads, which decreased to 96 after 30 min, 86 after 2h, 44 after 4h, and reached a minimum of 29 after 24h (Fig. 2). Most of the TUGs encode for chlorophyll a-b binding proteins that transfer energy between photosystem antennae and reaction centers, and suggests a general down-regulation of photosynthetic activity. This is consistent with physiological studies showing that the photosynthetic yield of *Z. noltii* decreases immediately after a heat-shock [5, 35]. The photosynthetic machinery is integrated and composed of many subunits making it energetically expensive for the cell to produce under stress [30]. Therefore, the down-regulation of TUGs related to photosynthetic components under heat-shock, such as highly abundant chlorophyll a-b binding proteins, could reveal a trade-off that helps maintain energy balance. However, since previous work has shown that about half of the plants will eventually die following a similar heat-shock [5], under-expression of photosynthesis TUGs may also be a consequence of apoptosis.

An exception to the general pattern in light-harvesting proteins was the up-regulation after 2 h of a CP24 protein (ELIP/psbS family; Figure 2). Members of this protein family are thought to be important in non-photochemical quenching of excess light energy and protection against photo-oxidative stress [36].

### *3.5. Other TUGs of interest*

A number of other TUGs with diverse functions were also found to have significantly different expressions among libraries (Fig. 3).

320 Ubiquitin

321 Ubiquitin is a highly conserved small protein, involved in the selective degradation of
322 many short-lived proteins in eukaryotic cells as they are targeted for degradation by
323 covalent ligation to ubiquitin. Ubiquitin-mediated degradation of regulatory proteins
324 plays important roles in the control of numerous processes, including cell-cycle
325 progression, signal transduction, transcriptional regulation, receptor down-regulation,
326 and endocytosis. The ubiquitin system has been implicated in the immune response,
327 development, and programmed cell death [37, 38]. Ubiquitin TUGs were over-
328 expressed in the stress libraries, suggesting that heat-stress may have damaged some
329 proteins, which were targeted for proteolysis.

330

331 Glutamine synthetase

332 Another TUG that was also slightly elevated in the stress libraries was glutamine
333 synthetase (GS), involved in the assimilation of nitrogen and a biomarker of plant
334 metabolism, indicating nutrient deficiency under stress conditions [39]. Higher plant GS
335 in roots functions in the primary assimilation of ammonia from the soil. In leaves, GS is
336 also responsible for the reassimilation and detoxification of the large amounts of
337 ammonia lost during photorespiration [40]. Accumulation of nitrogen may be explained
338 by the lack of use of this nutrient for growth during stress conditions, as was observed
339 in *Z. marina* during reduced light conditions [41].

340

341 Response to oxidative stress and/or cellular detoxification

342 The classical plant peroxidases are a well-studied group of heme-containing enzymes
343 that utilize either $H_2O_2$ or $O_2$ to oxidize a wide variety of substrates [42]. In the majority
344 of plant species investigated they occur as distinctive isoenzymes which can be

345 constitutive or induced in response to external factors such as wounding, stress and
346 attack by pathogens [43]. Peroxidase showed an over-expression in the middle of the
347 heat-shock (2 hours) when compared to the remaining libraries.
348 The glutathione S-transferase (*GST*) super-family of genes encode enzymes that
349 catalyse a number of distinct glutathione-dependent reactions: in addition to their ability
350 to catalyse the formation of conjugates, *GST* can also serve as peroxidases and
351 isomerases. They play a critical role in protecting against electrophiles and products of
352 oxidative stress and cellular detoxification generally, suggesting that they are part of an
353 adaptive response to chemical stress [44]. This TUG was also over-expressed in the
354 stress libraries, suggesting that GST may also participate in the response to temperature
355 stress.
356 Plant metallothioneins are involved in heavy metal tolerance and detoxification
357 and might also be involved in regulation of cellular availability of required heavy
358 metals, namely copper and zinc [45, 46] and were under-expressed in the stress
359 libraries, suggesting cell energy may be switched from heavy metal detoxification to
360 heat stress response .
361
362 Reticulon-like protein
363 Reticulons are proteins that have been found predominantly associated with the
364 endoplasmic reticulum in yeast and mammalian cells. Although reticulon-like proteins
365 have been identified in plants, very little is known about their cellular localization and
366 functions, but have been shown to be associated with the endoplasmic reticulum in
367 *Arabidopsis thaliana* [47]. Reticulon-like proteins showed a slight variation among
368 stress libraries, being under-expressed during the beginning of the heat-shock (30

minutes and 2 hours libraries), suggesting a decrease in a hypothetical housekeeping gene during high stress conditions.

Peptidyl-prolyl isomerase

Peptidyl prolyl isomerases are protein folding catalysts [48, 49], whose function is the cis-trans isomerization of peptidyl-prolyl bonds, a relevant conformational change that is rate limiting. Most, but not all, peptide bonds are connected in the trans conformation during biosynthesis at the ribosomes, and this conformation is also found in the native structure of most peptide bonds [50]. The over-expression of these TUGs in the stress libraries may suggest an increase in protein biosynthesis, which is possibly related to the production of stress-response molecules such as heat-shock proteins.

Isoflavone reductase homolog

Isoflavone reductase (*IFR*) is an enzyme specific to isoflavonoid biosynthesis, a pathway which is mainly found in the *Leguminosae* (angiosperms). It catalyses a NADPH-dependent reduction involved in the biosynthesis of important and related phenylpropanoid-derived plant defense compounds [51]. Interestingly, this enzyme was over-expressed in the 24 hours library, which may reflect a delayed response to the stress.

Flavoprotein WrbA

The flavoprotein WrbA, originally described as a tryptophan (*W*) *r*epressor-*b*inding protein in *Escherichia coli*, has recently been shown to exhibit the enzymatic activity of a NADH: quinone oxidoreductase to maintain of a supply of reduced quinine, and having a possible role in stress response [52]. WrBA was under-expressed in the first

stress library (30 minutes), which suggests this enzyme might be inhibited by high temperatures.

Glyceraldehyde-3-phosphate dehydrogenase

Glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) is a nucleic-acid-binding protein involved in glycolysis and in the Calvin cycle [53-55] and was under-expressed under stress conditions, suggesting an inhibition of important metabolic pathways during high temperature stress. *GAPDH* exhibits high temperature sensitivity [56] and the inhibition of this enzyme affects the photosynthetic process, by causing a secondary photoinhibitory response in PSII [57] that may be the cause of the decrease in photosynthetic activity. It can also reduce energy generation if glycolysis is inhibited and ATP levels decrease.

### *3.6. Microsatellite identification*

A total of 4,734 TUGs were screened for EST-SSRs and 86 microsatellites were identified, with 43 di-, 36 tri-, 4 tetra- and 3 hexanucleotides (Table 3). The main advantage of EST-SSR over SSR markers developed by random genomic screening is their physical linkage to coding sequences. In future studies, these may be used to identify genomic regions associated with heat stress response, as markers can be designed around a SSR in a candidate gene. Another advantage is their apparently high conservation between species and even genera, probably because these are more likely to be in gene-rich regions of chromosomes, which can facilitate the screening of related species even when no EST resources are available [58].

## 4. Conclusions

This study reports the first transcriptomic dataset for *Z. noltii,* focusing on response to high temperature stress. Understanding the molecular basis of traits of interest has been hindered by a lack of genomic resources for this species, and this study has provided a considerable dataset covering the transcriptional response to heat-stress. Almost 8,600 sequences reads were produced from all libraries, which resulted in over 3,000 annotated TUGs. As expected, an important part of the TUGs with significantly different expressions between libraries were known to be stress-related, mostly heat-shock proteins, which were over-expressed in stress libraries and are commonly induced by high temperature stress. Also, photosynthetic activity seems to be affected by heat-shock, as important enzymes are thermally inhibited and cells redirect their energy towards defense strategies. The heat-shock response in *Z. noltii* is very fast, showing significant differences from unstressed plants as soon as 30 minutes after the beginning of the stress, causing an increase of heat-shock protein TUGs and an inhibition of some of the TUGs with photosynthetic function. A total of 86 microsatellites were also identified, and in future work may be used to develop genetic markers in candidates TUGs related to the response to heat stress.

**5. Acknowledgments:**

This research was funded by the project DIVSTAB (POCI/MAR/60179/2004) of the Portuguese Science Foundation FCT, FEDER, the Marine Genomics Network of Excellence (MGE GOCE-CT-2003-505446) and a fellowship from FCT and ESF (to SM). We thank Sven Klages and Cymon Cox for their assistance in the data analysis and João Reis and Miguel Viegas for their technical assistance in the setup of experimental designs.

## 6. References


[1] R.J. Orth, T.J.B. Carruthers, W.C. Dennison, C.M. Duarte, J.W. Fourqurean, K.L. Heck, A.R. Hughes, G.A. Kendrick, W.J. Kenworthy, S. Olyarnik, F.T. Short, M. Waycott, S.L. Williams, A global crisis for seagrass ecosystems, Bioscience, 56 (2006) 987-996.

[2] M. Waycott, C.M. Duarte, T.J.B. Carruthers, R.J. Orth, W.C. Dennison, S. Olyarnik, A. Calladine, J.W. Fourqurean, K.L. Heck, A.R. Hughes, G.A. Kendrick, W.J. Kenworthy, F.T. Short, S.L. Williams, Accelerating loss of seagrasses across the globe threatens coastal ecosystems, Proc. Natl. Acad. Sci. U. S. A., 106 (2009) 12377-12381.

[3] J.E. Duffy, Biodiversity and the functioning of seagrass ecosystems, Mar. Ecol. Prog. Ser., 311 (2006) 233-250.

[4] IPCC, Climate Change 2007: The Physical Science Basis. Summary for Policymakers., (2007) 1-21.

[5] S.I. Massa, S. Arnaud-Haond, G.A. Pearson, E.A. Serrao, Temperature tolerance and survival of intertidal populations of the seagrass Zostera noltii (Hornemann) in Southern Europe (Ria Formosa, Portugal), Hydrobiologia, 619 (2009) 195-201.

[6] R. Arya, M. Mallik, S.C. Lakhotia, Heat shock genes - integrating cell survival and death, Journal of Biosciences, 32 (2007) 595-610.

[7] M.J. Schlesinger, Heat-Shock Proteins, J. Biol. Chem., 265 (1990) 12111-12114.

[8] S.J. Campbell, L.J. McKenzie, S.P. Kerville, Photosynthetic responses of seven tropical seagrasses to elevated seawater temperature, J. Exp. Mar. Biol. Ecol., 330 (2006) 455-468.

[9] L.J. Cui, J.L. Li, Y.M. Fan, S. Xu, Z. Zhang, High temperature effects on photosynthesis, PSII functionality and antioxidant activity of two Festuca arundinacea cultivars with different heat susceptibility, Botanical Studies, 47 (2006) 61-69.

[10] M. Almeselmani, P.S. Deshmukh, S. R.K., S.R. Kushwaha, T.P. Singh, Protective role of antioxidant enzymes under high temperature stress, Plant Science, 171 (2006) 382-388.

[11] J.J. Bull, M.R. Badgett, H.A. Wichman, J.P. Huelsenbeck, D.M. Hillis, A. Gulati, C. Ho, I.J. Molineux, Exceptional convergent evolution in a virus, Genetics, 147 (1997) 1497-1507.

[12] R. Woods, D. Schneider, C.L. Winkworth, M.A. Riley, R.E. Lenski, Tests of parallel molecular evolution in a long-term experiment with Escherichia coli, Proc. Natl. Acad. Sci. U. S. A., 103 (2006) 9107-9112.

[13] J. de Visser, A.D.L. Akkermans, R.F. Hoekstra, W.M. de Vos, Insertion-sequence-mediated mutations isolated during adaptation to growth and starvation in Lactococcus lactis, Genetics, 168 (2004) 1145-1157.

[14] S.E. Travers, Z.W. Tang, D. Caragea, K.A. Garrett, S.H. Hulbert, J.E. Leach, J.F. Bai, A. Saleh, A.K. Knapp, P.A. Fay, J. Nippert, P.S. Schnable, M.D. Smith, Variation in gene expression of Andropogon gerardii in response to altered environmental conditions associated with climate change, J. Ecol., 98 (2010) 374-383.

[15] H.H. Chou, M.H. Holmes, DNA sequence quality trimming and vector removal, Bioinformatics, 17 (2001) 1093-1104.

[16] G. Pertea, X.Q. Huang, F. Liang, V. Antonescu, R. Sultana, S. Karamycheva, Y. Lee, J. White, F. Cheung, B. Parvizi, J. Tsai, J. Quackenbush, TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets, Bioinformatics, 19 (2003) 651-652.

[17] A. Conesa, S. Gotz, J.M. Garcia-Gomez, J. Terol, M. Talon, M. Robles, Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research, Bioinformatics, 21 (2005) 3674-3676.
[18] S. Gotz, J.M. Garcia-Gomez, J. Terol, T.D. Williams, S.H. Nagaraj, M.J. Nueda, M. Robles, M. Talon, J. Dopazo, A. Conesa, High-throughput functional annotation and data mining with the Blast2GO suite, Nucleic Acids Res., 36 (2008) 3420-3435.
[19] C. Romualdi, S. Bortoluzzi, F. D'Alessi, G.A. Danieli, IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments, Physiol. Genomics, 12 (2003) 159-162.
[20] S. Audic, J.M. Claverie, The significance of digital gene expression profiles, Genome Res., 7 (1997) 986-995.
[21] T.B.H. Reusch, A.S. Veron, C. Preuss, J. Weiner, L. Wissler, A. Beck, S. Klages, M. Kube, R. Reinhardt, E. Bornberg-Bauer, Comparative analysis of expressed sequence tag (EST) libraries in the seagrass Zostera marina subjected to temperature stress, Mar. Biotechnol., 10 (2008) 297-309.
[22] L. Wissler, E. Dattolo, A.D. Moore, T.B.H. Reusch, J.L. Olsen, M. Migliaccio, E. Bornberg-Bauer, G. Procaccini, Dr. Zompo: an online data repository for *Zostera marina* and *Posidonia oceanica* ESTs, Database, (2009).
[23] B.C. Faircloth, MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design, Molecular Ecology Resources, 8 (2008) 92-94.
[24] X.J. Min, G. Butler, R. Storms, A. Tsang, OrfPredictor: predicting protein-coding regions in EST-derived sequences, Nucleic Acids Res., 33 (2005) W677-W680.
[25] P. Meinicke, UFO: a web server for ultra-fast functional profiling of whole genome protein sequences, BMC Genomics, 10 (2009).
[26] P.S. Krishna, Plant responses to heat stress, in: H. Hirt, K. Shinozaki (Eds.) Plant responses to abiotic stress, Springer-Verlag, Berlin Heidelberg, 2004, pp. 9-38.
[27] H.M. Beere, Death versus survival: functional interaction between the apoptotic and stress-inducible heat shock protein pathways, J. Clin. Investig., 115 (2005) 2633-2639.
[28] M.E. Cheetham, A.J. Caplan, Structure, function and evolution of DnaJ: conservation and adaptation of chaperone function, Cell Stress Chaperones, 3 (1998) 28-36.
[29] D.M. Cyr, T. Langer, M.G. Douglas, Dnaj-Like Proteins - Molecular Chaperones and Specific Regulators of Hsp70, Trends Biochem. Sci., 19 (1994) 176-181.
[30] G. Govind, H.V. ThammeGowda, P.J. Kalaiarasi, D.R. Iyer, S.K. Muthappa, S. Nese, U.K. Makarla, Identification and functional validation of a unique set of drought induced genes preferentially expressed in response to gradual water stress in peanut, Molecular Genetics and Genomics, 281 (2009) 591-605.
[31] S.G. Ahn, S.A. Kim, J.H. Yoon, P. Vacratsis, Heat-shock cognate 70 is required for the activation of heat-shock factor 1 in mammalian cells, Biochem. J., 392 (2005) 145-152.
[32] P. Genevaux, C. Georgopoulos, W.L. Kelley, The Hsp70 chaperone machines of Escherichia coli: a paradigm for the repartition of chaperone functions, Mol. Microbiol., 66 (2007) 840-857.
[33] K. Aoki, F. Kragler, B. Xoconostle-Cazares, W.J. Lucas, A subclass of plant heat shock cognate 70 chaperones carries a motif that facilitates trafficking through plasmodesmata, Proc. Natl. Acad. Sci. U. S. A., 99 (2002) 16342-16347.
[34] C.M. Nicolet, E.A. Craig, Isolation and Characterization of Sti1, a Stress-Inducible Gene from Saccharomyces-Cerevisiae, Mol. Cell. Biol., 9 (1989) 3638-3646.

542 [35] P.J. Ralph, Photosynthetic response of laboratory-cultured Halophila ovalis to
543 thermal stress, Mar. Ecol. Prog. Ser., 171 (1998) 123-130.
544 [36] C. Hutin, L. Nussaume, N. Moise, I. Moya, K. Kloppstech, M. Havaux, Early light-
545 induced proteins protect arabidopsis from photooxidative stress, Proc. Natl. Acad. Sci.
546 U. S. A., 100 (2003) 4921-4926.
547 [37] E.S. Johnson, B. Bartel, W. Seufert, A. Varshavsky, Ubiquitin as a Degradation
548 Signal, Embo Journal, 11 (1992) 497-505.
549 [38] A. Hershko, A. Ciechanover, The ubiquitin system, Annu. Rev. Biochem., 67
550 (1998) 425-479.
551 [39] L. Ferrat, C. Pergent-Martini, M. Romeo, Assessment of the use of biomarkers in
552 aquatic plants for the evaluation of environmental quality: application to seagrasses,
553 Aquatic Toxicology, 65 (2003) 187-204.
554 [40] A.J. Keys, I.F. Bird, M.J. Cornelius, P.J. Lea, R.M. Wallsgrove, B.J. Miflin,
555 Photorespiratory Nitrogen Cycle, Nature, 275 (1978) 741-743.
556 [41] F. Vanlent, J.M. Verschuure, M.L.J. Vanveghel, Comparative-Study on
557 Populations of Zostera-Marina L (Eelgrass) - in-Situ Nitrogen Enrichment and Light
558 Manipulation, J. Exp. Mar. Biol. Ecol., 185 (1995) 55-76.
559 [42] K. Yoshida, P. Kaothien, T. Matsui, A. Kawaoka, A. Shinmyo, Molecular biology
560 and application of plant peroxidase genes, Appl. Microbiol. Biotechnol., 60 (2003) 665-
561 670.
562 [43] N.C. Veitch, Horseradish peroxidase: a modern view of a classic enzyme,
563 Phytochemistry, 65 (2004) 249-259.
564 [44] J.D. Hayes, D.J. Pulford, The glutathione S-Transferase supergene family:
565 Regulation of GST and the contribution of the isoenzymes to cancer chemoprotection
566 and drug resistance, Crit. Rev. Biochem. Mol. Biol., 30 (1995) 445-600.
567 [45] N.J. Robinson, A.M. Tommey, C. Kuske, P.J. Jackson, Plant Metallothioneins,
568 Biochem. J., 295 (1993) 1-10.
569 [46] S. Moisyadi, J.I. Stiles, A Cdna-Encoding a Metallothionein I-Like Protein from
570 Coffee Leaves (Coffea-Arabica), Plant Physiology, 107 (1995) 295-296.
571 [47] H. Nziengui, K. Bouhidel, D. Pillon, C. Der, F. Marty, B. Schoefs, Reticulon-like
572 proteins in Arabidopsis thaliana: Structural organization and ER localization, FEBS
573 Lett., 581 (2007) 3356-3362.
574 [48] F.X. Schmid, Protein-Folding - Prolyl Isomerases Join the Fold, Curr. Biol., 5
575 (1995) 993-994.
576 [49] M. Kruse, M. Brunke, A. Escher, A.A. Szalay, M. Tropschug, R. Zimmermann,
577 Enzyme Assembly after De-Novo Synthesis in Rabbit Reticulocyte Lysate Involves
578 Molecular Chaperones and Immunophilins, J. Biol. Chem., 270 (1995) 2588-2594.
579 [50] S.F. Gothel, M.A. Marahiel, Peptidyl-prolyl cis-trans isomerases, a superfamily of
580 ubiquitous folding catalysts, Cellular and Molecular Life Sciences, 55 (1999) 423-436.
581 [51] S.C. Franca, P.G. Roberto, M.A. Marins, R.D. Puga, A. Rodrigues, J.O. Pereira,
582 Biosynthesis of secondary metabolites in sugarcane, Genet. Mol. Biol., 24 (2001) 243-
583 250.
584 [52] S.L.A. Andrade, E.V. Patridge, J.G. Ferry, O. Einsle, Crystal structure of the
585 NADH : Quinone oxidoreductase WrbA from Escherichia coli, J. Bacteriol., 189 (2007)
586 9101-9107.
587 [53] R.D. Barber, D.W. Harmer, R.A. Coleman, B.J. Clark, GAPDH as a housekeeping
588 gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues, Physiol.
589 Genomics, 21 (2005) 389-395.
590 [54] N. Wedel, J. Soll, Evolutionary conserved light regulation of Calvin cycle activity
591 by NADPH-mediated reversible phosphoribulokinase/CP12/glyceraldehyde-3-

phosphate dehydrogenase complex dissociation, Proc. Natl. Acad. Sci. U. S. A., 95 (1998) 9699-9704.
[55] W.C. Plaxton, The organization and regulation of plant glycolysis, Annu. Rev. Plant Physiol. Plant Mol. Biol., 47 (1996) 185-214.
[56] T. Mair, C. Warnke, K. Tsuji, S.C. Muller, Control of glycolytic oscillations by temperature, Biophys. J., 88 (2005) 639-646.
[57] P.J. Ralph, Photosynthetic response of Halophila ovalis (R-Br.) Hook. f. to combined environmental stress, Aquat. Bot., 65 (1999) 83-96.
[58] A. Bouck, T. Vision, The molecular ecologist's guide to expressed sequence tags, Mol. Ecol., 16 (2007) 907-924.

606 **Table 1: Summary of the sequence read analysis for each heat-shock library (30**

607 **min, 2h, 4h and 24h) and the control (blank)**.

| **Library** | **30 min** | **2h** | **4h** | **24h** | **control** |
|---|---|---|---|---|---|
| **Total number of clones sequenced** | 1,724 | 1,795 | 1,756 | 1,734 | 1,585 |
| **Number of high quality sequences** | 1,657 (96.1%) | 1,698 (94.6%) | 1,553 (88.4%) | 1,563 (90.1%) | 1,328 (83.8%) |
| **Total number of successful sequences** | 1,635 (94.8%) | 1,671 (93.1%) | 1,536 (87.5%) | 1,535 (88.5%) | 1,309 (82.6%) |
| **Number of contigs**[a] | 170 | 174 | 151 | 145 | 145 |
| **Number of clones included in the contigs** | 733 (44.8%) | 700 (41.9%) | 479 (31.2%) | 439 (28.6%) | 511 (39.0%) |
| **Average contig length** | 755 bp | 783 bp | 737 bp | 736 bp | 682 bp |
| **Number of singletons**[b] | 902 (55.2%) | 971 (58.1%) | 1,057 (68.8%) | 1,096 (71.4%) | 798 (61.0%) |
| **Tentative UniGenes (TUGs)** [(a+b)] | 1,072 | 1,145 | 1,208 | 1,241 | 943 |
| **Number of TUG with no ORF** | 2 (0.18%) | 4 (0.35%) | 1 (0.08%) | 1 (0.08%) | 13 (1.38%) |
| **Number of TUGs with a successful blast** | 907 (84.6%) | 985 (86.0%) | 1,030 (85.3%) | 1,031 (83.1%) | 690 (73.2%) |
| **Number of successfully mapped TUGs** | 745 (69.5%) | 824 (72.0%) | 832 (68.9%) | 914 (73.7%) | 628 (66.6%) |
| **Number of successfully annotated TUGs** | 600 (56.0%) | 639 (55.8%) | 650 (53.8%) | 730 (58.8%) | 527 (55.9%) |

608

609 **Table 2: Main protein families.** All families with at least 10 sequence reads in one of

610 the libraries.

| | 30 min | 2h | 4h | 24h | control |
|---|---|---|---|---|---|
| (Hsp70 protein) | 28 | 31 | 16 | 13 | 1 |
| (Protein kinase domain) | 16 | 23 | 19 | 20 | 12 |
| (Chlorophyll A-B binding protein) | 16 | 15 | 16 | 15 | 16 |
| (RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)) | 11 | 8 | 8 | 8 | 5 |
| (Cytochrome P450) | 6 | 2 | 10 | 2 | 4 |

611

612 **Table 3: EST-SSRs found in all TUGs.**

| | Number of repeat units | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Motif** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **Total** |
| AC | | 2 | 1 | | | | | | | | | 3 |
| AG | 6 | 2 | 2 | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | 17 |
| AT | 8 | 7 | 2 | 1 | 1 | 2 | | 1 | | 1 | | 23 |
| AAC | 1 | | | | | | | | | | | 1 |
| AAG | 7 | 8 | 1 | 1 | | | 1 | 1 | | | | 19 |
| AAT | 4 | | | | | | | | | | | 4 |
| ACC | | 1 | 1 | | | | | | | | | 2 |
| ACT | 1 | | | | | | | | | | | 1 |
| AGC | 1 | | | | | | | | | | | 1 |
| AGG | 2 | 2 | | | | | | | | | | 4 |
| ATC | 1 | 1 | | | | | | | | | | 2 |
| CCG | 2 | | | | | | | | | | | 2 |
| AATG | | | | | 1 | | | | | | | 1 |
| AATT | | 1 | | | | | | | | | | 1 |
| AGAT | 1 | | | | | | | | | | | 1 |
| ATCC | | 1 | | | | | | | | | | 1 |
| AACATG | 1 | | | | | | | | | | | 1 |
| AATCAC | 1 | | | | | | | | | | | 1 |
| AGATGG | 1 | | | | | | | | | | | 1 |
| Total | 31 | 19 | 7 | 3 | 1 | 1 | 2 | 3 | 1 | 2 | 1 | 86 |

613

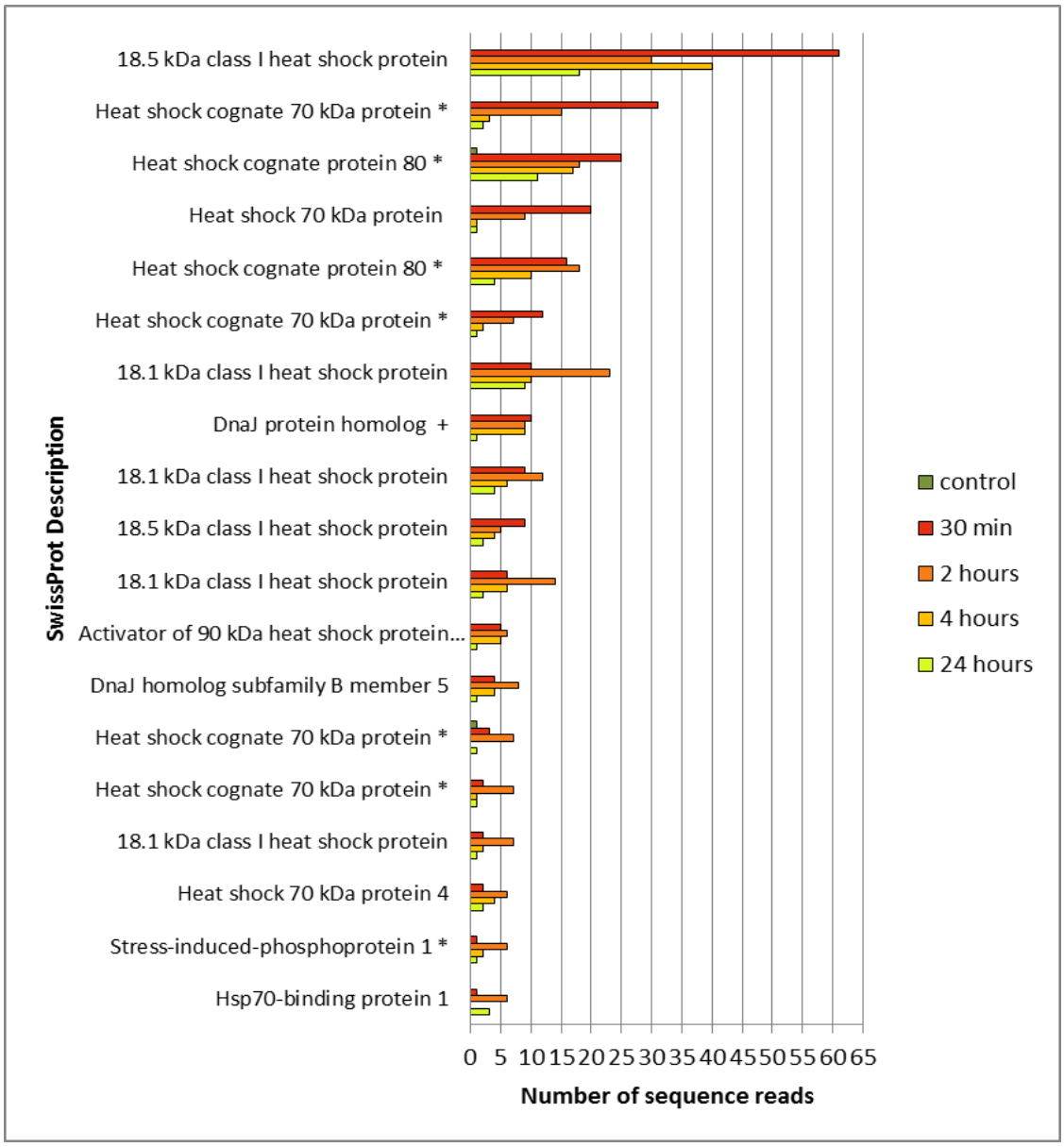
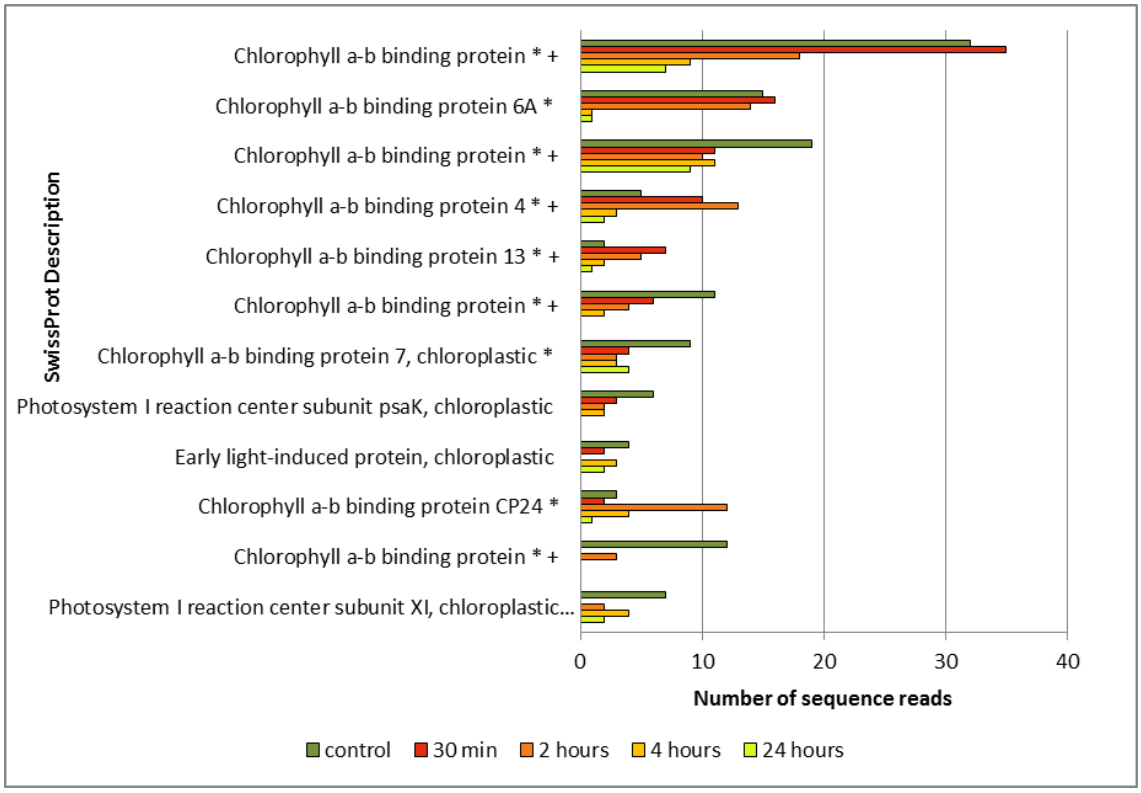



**Figure 1: Number of sequence reads of each 'molecular chaperone' TUG with significantly different expressions in each library.** SwissProt descriptions were used to identify each TUG. The number of sequence reads for the control library is shown in white, stress libraries are as follows: 30 minutes in black, 2h in dark grey, 4h in medium grey and 24h in light grey. It is also shown when a corresponding hit was found in Dr. Zompo database for *Zostera marina* (*) and *Posidonia oceanica* (+).




621

622 **Figure 2: Number of sequence reads of each photosynthesis-related TUG with**
623 **significantly different expressions in each library.** SwissProt descriptions were used
624 to identify each TUG. The number of sequence reads for the control library is shown in
625 white, stress libraries are as follows: 30 minutes in black, 2h in dark grey, 4h in medium
626 grey and 24h in light grey. It is also shown when a corresponding hit was found in Dr.
627 Zompo database for *Zostera marina* (*) and *Posidonia oceanica* (+).

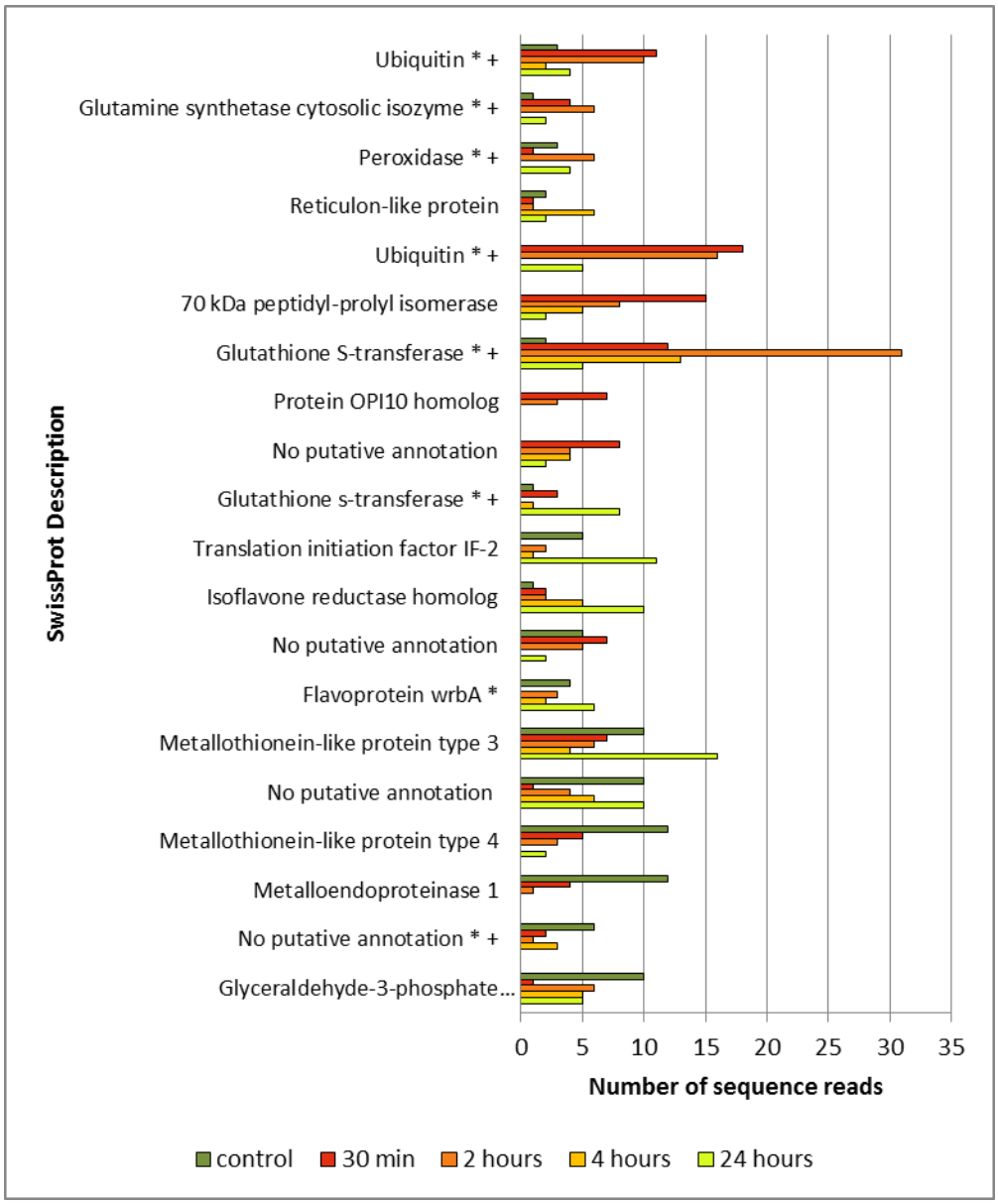



**Figure 3: Number of sequence reads of other TUGs of interest with significantly different expressions in each library.** SwissProt descriptions were used to identify each TUG. The number of sequence reads for the control library is shown in white, stress libraries are as follows: 30 minutes in black, 2h in dark grey, 4h in medium grey and 24h in light grey. It is also shown when a corresponding hit was found in Dr. Zompo database for *Zostera marina* (*) and *Posidonia oceanica* (+).